Dear AD,

I hope this email finds you well. I have thoroughly considered the hypothesis regarding churn being driven by customers' price sensitivities, and I have formulated a plan to test this hypothesis. Below, I outline the major steps required along with the necessary data and analytical models to support our investigation.

Data Collection:

To test the hypothesis, we would need access to the following data from PowerCo:

a. **Customer Data**: This includes information such as customer demographics, type (corporate, SME, residential), contract start and end dates, historical energy consumption, and any other relevant attributes that can provide insights into customer behavior.

b. **Billing Data**: This includes details of the energy bills sent to customers, including the billed amount, payment history, and any discounts or special offers applied.

c. **Churn Data**: We need a record of churn events, specifically customers who have terminated their contract with PowerCo within a certain period. This data should ideally include the churn date, reasons for churn (if available), and any actions taken to prevent churn.

d. Competitor Pricing Data: It would be valuable to obtain data on competitor pricing in the market to compare it with Power Co's pricing. This can help us understand how competitive Power Co's prices are and their potential impact on customer churn.

Data Preparation:

Once we have the necessary data, we need to clean and preprocess it to ensure it is suitable for analysis. This involves handling missing values, formatting inconsistencies, and merging relevant datasets. We should also create a unified data frame where each row represents a customer, and the columns contain the relevant customer attributes and churn status.

Exploratory Data Analysis (EDA):

Performing EDA on the collected data will provide valuable insights into the relationship between price sensitivity and churn. Here are a few analyses we can conduct:

a. Churn Rate Analysis: Calculate the churn rate for different customer segments (corporate, SME, residential) and compare it against their price sensitivity based on historical pricing data. This will help identify any patterns or correlations between price sensitivity and churn.

b. Price Distribution Analysis: Examine the distribution of prices charged to customers and identify any trends or clusters that may be associated with churn behavior.

c. Customer Segmentation: Cluster customers based on their attributes and behavior using techniques like clustering algorithms (e.g., K-means) or hierarchical clustering. This can reveal distinct customer segments with different price sensitivities and churn probabilities.

d. Correlation Analysis: Explore correlations between churn and other customer attributes, such as contract duration, payment history, energy consumption patterns, and demographic factors. This will help identify additional factors that contribute to churn, beyond price sensitivity.

Predictive Modeling:

To predict which customers are more likely to churn at their current price, we can develop a predictive model using techniques such as logistic regression, decision trees, or random forests. The model should utilize the customer attributes as input features and the churn status as the target variable. By training the model on historical data, we can assess its accuracy in predicting churn and identify the key features driving customer churn.

Model Evaluation and Validation:

Evaluate the predictive model's performance using appropriate evaluation metrics such as accuracy, precision, recall, and F1-score. Validate the model using a holdout dataset or cross-validation techniques to ensure its generalizability.

Discount Strategy:

Once the predictive model is validated, we can use it to identify customers who are at risk of churning. For those customers, we can recommend offering a 20% discount to dissuade them from switching to competitors. However, it's essential to track the effectiveness of the discount strategy by monitoring churn rates and customer retention over time.

In conclusion, by following these steps and utilizing the available data, we can thoroughly test the hypothesis that churn in the SME segment is driven by customers' price sensitivities. The insights gained from the analysis will help PowerCo make informed decisions regarding discount offers and customer retention strategies.

Please let me know if you have any further questions or if there are any additional aspects you would like me to consider. I look forward to discussing this further with you.

Best regards,

Sandeep Yempaty.